AJSTME, Volume. 12, Issue 1. 304-313, January, 2026. https://www.ajstme.com.ng
ISSN: 2251-0141 (Print), 2971-6233 (Online)

An Evaluation of the Effectiveness of AI-Based Learning Tools on Cognitive Achievement in Educational Statistics

Onyemachi, Chukwudi Godwin[1] & Frederick Ekene Onah[2]
[1]College of Nursing Sciences Emekuku Owerri
[2]Faculty of Education, Imo State University Owerri
Corresponding Author: onyemachichukwudigodwin@gmail.com

Abstract

Educational Statistics remains a challenging course for many undergraduates, often resulting in low cognitive achievement due to abstract concepts and limited individualized support. This study examined the effectiveness of Artificial Intelligence (AI)-based learning tools on students' achievement in Educational Statistics. A quasi-experimental non-equivalent control group design was adopted with 200 students (100 AI-exposed, 100 non-AI). Data were collected using the Cognitive Achievement Test in Educational Statistics (CATES). Mean and standard deviation answered research questions, while ANCOVA tested hypotheses at 0.05 significance level. Results showed that AI-exposed students achieved higher post-test scores ($M = 32.96$, $SD = 4.88$) than non-AI students ($M = 27.14$, $SD = 5.02$), with a significant difference, $F(1,196) = 27.89$, $p < .001$. No significant gender differences were found. The study concludes that AI tools significantly improve achievement and recommends their integration into statistics instruction.

Keywords: Adaptive Learning, Artificial Intelligence, Classroom Assessment Performance, Cognitive Achievement, Educational Statistics, Undergraduate Students

Introduction

The rapid advancement of Artificial Intelligence (AI) in recent years has transformed many sectors globally, including education. AI-driven learning tools such as intelligent tutoring systems, adaptive learning platforms, automated feedback systems, and AI-supported assessment tools are increasingly being integrated into higher education to enhance students' learning experiences and outcomes (Okonkwo & Ahmed, 2023; UNESCO, 2022). In the domain of Educational Statistics, which requires analytical thinking, problem-solving skills, and continuous practice, AI-based learning tools have emerged as innovative mechanisms that provide personalized feedback, adaptive difficulty levels, and interactive learning environments that traditional lecture-based teaching cannot fully achieve.

Cognitive achievement, which refers to the intellectual skills students acquire in the learning process, is a key measure of academic success, particularly in quantitative courses such as statistics (Anderson & Krathwohl, 2021). However, students offering Educational Statistics often encounter difficulties in understanding statistical concepts, interpreting results, and applying statistical procedures challenges that contribute to poor performance in both theoretical and practical components (Nworgu & Ugwuanyi, 2022). These challenges necessitate alternative instructional strategies, including the integration of AI-driven learning innovations, to promote deeper understanding and improve classroom assessment performance.

AI-based tools such as ChatGPT, machine learning tutors, adaptive statistical simulators, AI-powered formative assessment platforms, and automated grading systems are designed to support individualized learning by diagnosing students' weaknesses and offering tailored instructional support (Holmes et al., 2021). This aligns with contemporary pedagogical approaches that emphasize learner-centered education, constructivist learning, and technology-enhanced instruction.

In Nigeria, the adoption of AI-driven educational tools is growing, particularly within higher institutions striving to meet global standards in digital learning (Adebayo & Salami, 2023). Faculties of Education, which are mandated to train future teachers with modern pedagogical competencies, play a critical role in modeling innovative instructional practices. Undergraduate students offering Educational Statistics in the Faculty of Education represent a unique group that often exhibits varied cognitive capabilities, diverse learning needs, and inconsistent exposure to digital tools. Consequently, investigating the effectiveness of AI-based learning tools on their cognitive achievement and classroom assessment performance is timely, relevant, and academically significant.

Despite the potential of AI tools to improve learning outcomes, empirical evidence within Nigerian universities especially regarding Educational Statistics remains limited. This study is therefore positioned to fill a critical gap by evaluating how AI-based learning tools contribute to undergraduate students' mastery of statistical concepts, problem-solving skills, and overall classroom assessment performance.

In many Faculties of Education across Nigeria, undergraduate students offering Educational Statistics consistently exhibit low cognitive achievement and poor performance in both formative and summative classroom assessments (Nwafor & Onah, 2022). These performance gaps are attributed to several factors, including inadequate instructional materials, insufficient practice opportunities, teacher-centered pedagogies, and students' anxiety towards quantitative courses. Traditional teaching approaches, which rely heavily on lectures and manual exercises, often fail to accommodate the diverse learning needs of students. As a result, many learners struggle with complex statistical concepts such as probability distributions, hypothesis testing, regression models, and data interpretation.

AI-based learning tools hold promise for addressing these challenges by offering personalized learning pathways, instant feedback, automated correction of misconceptions, and interactive simulations. However, while these tools are increasingly available, their actual impact on cognitive learning outcomes and classroom assessment performance among undergraduate students remains unclear within the Nigerian university context. Specifically, at the Faculty of Education, there is inadequate empirical data on the extent to which AI-assisted learning enhances students' understanding of Educational Statistics.

Artificial Intelligence (AI) refers to computer systems that can perform tasks that typically require human intelligence, such as problem-solving, reasoning, pattern recognition, and decision-making (Russell & Norvig, 2021). In the context of education, AI enhances learning by enabling adaptive instruction, personalized learning paths, automated feedback, and intelligent assessment systems (Holmes et al., 2021). AI applications in education include chatbots, intelligent tutoring systems, virtual learning assistants, machine learning–driven analytics, and content recommendation systems.

In recent years, higher institutions globally have incorporated AI-based platforms such as ChatGPT, Knewton, Coursera AI, Google Classroom AI add-ons, and adaptive statistical simulators to improve teaching and learning outcomes. These tools do not merely supplement instruction but facilitate deeper conceptual understanding through real-time feedback, differentiated learning support, and immediate clarification of complex topics (UNESCO, 2022).

Moreover, concerns have been raised about students' over-dependence on AI tools, potential misuse, and disparities in digital literacy levels, which could affect learning outcomes. Without clear evidence of effectiveness, lecturers and curriculum planners lack the basis for making informed decisions regarding the integration of AI tools into Educational Statistics instruction.

Educational Statistics is a core course designed to equip students with knowledge of statistical methods for educational measurement, evaluation, and research. It includes topics such as descriptive statistics, probability distributions, correlation, regression, test development, and inferential statistics.

Adebayo and Salami (2023) found that Nigerian university students increasingly use AI-based tools for assignments, research, and problem solving, but utilization varies based on digital competence and internet accessibility. Similarly, Rahman and Putri (2022) reported positive attitudes toward AI tools among Indonesian undergraduates due to perceived usefulness in simplifying complex topics. However, a study by Chineke and Oliver (2022) revealed low awareness of AI-driven educational platforms among first-year students, suggesting that institution-wide exposure influences utilization.

Holmes et al. (2021) demonstrated that intelligent tutoring systems improved students' cognitive performance in mathematics and statistics more than traditional teaching alone. In another study, Ngugi and Eze (2020) found that chatbots significantly enhanced students' understanding of probability theory by providing immediate feedback. Uche and Okoro (2023) reported that students exposed to AI-powered simulations performed better in inferential statistics tests than those taught with traditional methods.

Nevertheless, some studies caution against over-dependence. Olanrewaju (2022) noted that while AI tools improved conceptual understanding, excessive reliance reduced students' independent problem-solving ability. According to Yuan and Chiang (2021), automated assessment tools improved consistency in grading and enhanced student performance by offering instant corrective feedback. Additionally, Mbata and Nwosu (2022) found that students who frequently engaged with AI practice tests scored significantly higher in final examinations. In contrast, Mensah (2021) observed that students with low digital literacy struggled to benefit from classroom AI-integrated AI-based learning tools have strong potential to improve cognitive achievement and classroom assessment performance, particularly in quantitative fields like Educational Statistics. Therefore, the problem of this study is the persistent low cognitive achievement and poor classroom assessment performance among undergraduate students offering Educational Statistics in the Faculty of Education, and the lack of empirical evidence on whether AI-based learning tools can effectively improve these outcomes. This study seeks to address this critical gap.

Research Questions

1. What are the mean achievement scores of students exposed to AI and those not exposed to AI tools?
2. What are the mean achievement scores of male and female students exposed to AI tools?
3. What are the mean achievement scores of male and female students not exposed to AI tools?

Hypotheses

$H0_1$: There is no significant difference in the mean achievement scores of students exposed to AI tools and those not exposed to AI tools.

$H0_2$: There is no significant difference in the mean achievement scores of male and female students exposed to AI tools.

$H0_3$: There is no significant difference in the mean achievement scores of male and female students not exposed to AI tools.

Methods

This study adopted a quasi-experimental research design, specifically the non-equivalent control group type, to examine the effect of Artificial Intelligence–based learning tools on students' achievement in Educational Statistics. The design was considered appropriate because intact classes were used without randomization, allowing the researcher to compare the performance of students exposed to AI-supported instruction with those taught through traditional lecture methods. The study was conducted in the Faculty of Education, Imo State University, Owerri, where Educational Statistics is offered as a compulsory course across several departments. This setting was suitable because it provided a heterogeneous group of learners with varying levels of digital exposure, which was necessary for distinguishing between AI-exposed and non-AI-exposed students.

The population of the study comprised all undergraduate students registered for Educational Statistics during the 2024/2025 academic session, totaling 1,340 students. From this population, a sample of 200 students was drawn through a multistage sampling process. First, students were stratified according to their academic levels from 100 to 400. Proportionate sampling was then applied to ensure that each level was represented in the same ratio as it existed within the overall population. Finally, purposive sampling was used to assign students into the two comparison groups. Students who regularly used AI-based learning tools such as ChatGPT, adaptive statistical simulators, and automated tutoring platforms were assigned to the AI-exposed group, while those who relied exclusively on traditional classroom instruction formed the non-AI group. Each group consisted of 100 students. The gender distribution of the sample comprised 90 males (45%) and 110 females (55%). Within the AI-exposed group, there were 46 males and 54 females, while the non-AI group consisted of 44 males and 56 females which helped to facilitate the comparison of male and female achievement scores.

Data were collected using three instruments. The primary instrument, the Cognitive Achievement Test in Educational Statistics (CATES), was a researcher-developed 40-item multiple-choice test that measured students' mastery of descriptive statistics, probability distributions, inferential statistics, correlation, regression, and hypotheses testing. A second instrument, the AI Learning Implementation Log (AILIL), was used to track the type, frequency, and duration of AI tool usage among students in the AI-exposed group. A third instrument, the Demographic and Background Data Sheet (DBDS), collected information on students' gender, level of study, and AI exposure status. The instruments were subjected to both content and construct validation by three specialists in Measurement and Evaluation, Statistics Education, and Educational Technology. Construct validation of the instruments was carried out using factor analysis procedures. The CATES items were subjected to exploratory factor analysis (EFA) to confirm that they adequately measured the intended constructs (descriptive statistics, inferential statistics, probability, and regression). Items with factor loadings below 0.40 were revised or discarded. Similarly, the AILIL was validated to ensure it accurately measured dimensions of AI usage such as frequency, duration, and type of interaction. The validation process confirmed that all retained items aligned with the theoretical constructs they were intended to measure. Item analysis was conducted on the CATES to improve its difficulty and discrimination indices, while its reliability was established using the Kuder–Richardson Formula 20, yielding a coefficient of 0.78. The AILIL demonstrated an internal consistency reliability coefficient of 0.82 using Cronbach's alpha. The study lasted for eight weeks. In the first week, the CATES was administered to both groups as a pre-test to establish equivalence in their baseline knowledge. Between the second and seventh weeks, the experimental group received instruction supplemented with AI-based learning tools. Students in this group interacted with AI systems that provided step-by-step explanations, adaptive learning pathways, automated corrections, and simulation-based practice opportunities. Their engagement with the AI tools was monitored using the AILIL. During the same period, the control group received instruction solely through teacher-led lectures, printed materials, and manual class exercises without the use of any AI tool. In the eighth week, the CATES was reshuffled and administered again as a post-test to both groups to measure cognitive achievement after the intervention. The reshuffling of test items was done to minimize test-retest bias and memorization effects, ensuring that students' post-test performance reflected actual learning rather than recall of item order. This approach enhances the internal validity of the study.

Data were analyzed using both descriptive and inferential statistics. The research questions were answered using mean and standard deviation to compare the achievement scores of students exposed to AI tools and those not exposed, as well as the gender-based comparisons within the two groups. The hypotheses were tested at the 0.05 level of significance. Analysis of Covariance (ANCOVA) was used to test the hypotheses comparing the AI-exposed and non-AI groups, with the pre-test scores serving as a covariate. The decision rule was that the null hypothesis would be rejected if the p-value is less than or equal to 0.05 ($p \leq 0.05$), and retained if the p-value is greater than 0.05 ($p > 0.05$). Ethical approval for the study was obtained from the university's research ethics committee, and all participants provided informed consent. Participation was voluntary, confidentiality was assured, and no student was disadvantaged by the intervention.

Control of Extraneous Variables

Extraneous variables were controlled using several strategies to enhance the internal validity of the study. First, pre-test scores were used as a covariate in the ANCOVA analysis to statistically control for initial differences in students' prior knowledge. Second, both groups were taught the same course content using a standardized instructional outline, ensuring uniformity in curriculum coverage. Third, the duration of instruction (eight weeks), learning environment, and assessment conditions were kept constant for both groups. In addition, efforts were made to minimize teacher-related variability by ensuring that instruction followed the same lesson structure and objectives across groups. To control for testing effects, the achievement test items were reshuffled during the post-test administration. Furthermore, instrument validity and reliability were established, reducing measurement error. These combined procedures helped to isolate the effect of AI-based learning tools on students' cognitive achievement.

Question 1: What are the mean achievement scores of students exposed to AI tools and those not exposed?

Table 1: Pre-test and Post-test Mean Achievement Scores of AI-Exposed and Non-AI Students

Group	N	Pre-test Mean	SD	Post-test Mean	SD	Mean Gain
AI-Exposed	100	21.84	4.91	32.96	4.88	11.12
Non-AI	100	22.03	4.87	27.14	5.02	5.11

The results in Table 1 show a clear distinction between the achievement outcomes of students exposed to AI-based learning tools and those who were not. The pre-test means (AI-exposed = 21.84; non-AI = 22.03) indicate that both groups were comparable at baseline, suggesting that the groups began the instructional period with essentially the same level of prior knowledge in Educational Statistics. This similarity strengthens the internal validity of the study because it implies that any differences observed in the post-test scores are attributable to the instructional treatment rather than pre-existing disparities.

After the instructional intervention, the AI-exposed group demonstrated a substantial improvement with a post-test mean of 32.96, compared to 27.14 for the non-AI group. The mean gain scores offer even clearer evidence of the differential impact of instruction: students taught with AI tools recorded a gain of 11.12 points, which is more than double the 5.11-point gain recorded by those taught using traditional methods. Such a large margin of improvement provides strong empirical support for the effectiveness of AI-supported learning environments in enhancing cognitive achievement in Educational Statistics.

The magnitude of improvement in the AI group suggests that AI tools played a vital role in facilitating deeper understanding. These tools provide real-time clarification, step-by-step explanations, adaptive difficulty adjustments, and immediate feedback—factors that are known to promote conceptual mastery. The relatively smaller improvement observed in the non-AI group reflects the limitations of traditional lecture-based teaching, which often provides delayed feedback, limited individualized support, and fewer opportunities for repeated diagnostic practice.

Overall, Table 1 supports the argument that AI-based instructional strategies provide a richer learning experience that leads to significantly higher learning gains than conventional instruction.

Question 2: What are the mean achievement scores of male and female students exposed to AI tools?

Table 2: Post-test Mean Achievement Scores of Male and Female Students Exposed to AI Tools

Gender	N	Pre-test Mean	SD	Post-test Mean	SD	Mean Gain
Male	46	22.01	4.85	33.41	4.72	11.40
Female	54	21.70	4.96	32.58	5.01	10.88

Table 2 presents important insights about gender differences within the AI-exposed group. The post-test means indicate that male students achieved a mean of 33.41, while female students achieved a mean of 32.58. Although males scored slightly higher, the difference of 0.83 points is educationally negligible. The similarity in performance demonstrates that AI-based learning tools provide a gender-neutral learning environment in which both male and female students benefit almost equally.

The standard deviations (males = 4.72; females = 5.01) further reinforce this observation, showing that the spread of scores is nearly identical across genders. This suggests that AI tools do not advantage one gender over the other in terms of variability or consistency of performance.

In essence, Table 2 highlights one of the key strengths of AI-enhanced learning: its capacity to level the playing field for all learners regardless of gender, thereby reducing gender-based achievement gaps commonly reported in quantitative courses.

Question 3: What are the mean achievement scores of male and female students not exposed to AI tools?

Table 3: Post-test Mean Achievement Scores of Male and Female Students Not Exposed to AI Tools

Gender	N	Pre-test Mean	SD	Post-test Mean	SD	Mean Gain
Male	44	22.15	4.89	27.70	5.13	5.55
Female	56	21.94	4.84	26.69	4.94	4.75

Table 3 presents the pre-test and post-test mean achievement scores of male and female students who were not exposed to AI-based learning tools. The results show that male students had a pre-test mean score of 22.15 (SD = 4.89), while female students recorded a slightly lower pre-test mean of 21.94 (SD = 4.84). The closeness of these pre-test mean scores indicates that both male and female students possessed comparable baseline knowledge in Educational Statistics prior to the instructional intervention. This similarity suggests that any subsequent differences in performance are unlikely to be attributed to initial ability differences. Following the instructional period, the post-test mean score for male students increased to 27.70 (SD = 5.13), while female students achieved a mean score of 26.69 (SD = 4.94). Although male students outperformed their female counterparts by 1.01 points, this difference is relatively small and educationally insignificant.

The mean gain scores provide further insight into learning progression. Male students recorded a mean gain of 5.55, while female students had a slightly lower gain of 4.75. These gains indicate that both groups benefited from the instructional process; however, the magnitude of improvement is moderate and notably lower than that observed in the AI-exposed group (Table 2). The standard deviation values for both male (SD = 5.13) and female (SD = 4.94) students are similar, suggesting comparable variability in performance within the two groups. This implies that the spread of scores and consistency of achievement are nearly the same across genders under traditional instruction.

Overall, the findings from Table 3 indicate that gender does not substantially influence students' achievement when AI tools are not used, as both male and female students' demonstrated similar levels of improvement and performance. However, the relatively lower post-test means and mean gains highlight the limitations of conventional lecture-based instruction, which appears less effective in enhancing cognitive achievement compared to AI-supported learning environments.

Hypotheses Testing

$H0_1$: There is no significant difference in the mean achievement scores of students exposed to AI tools and those not exposed.

To test this hypothesis, ANCOVA was used with the pre-test scores as covariate and post-test scores as dependent variable.

Table 4: ANCOVA Summary on Difference in Achievement between AI-Exposed and Non-AI Students

Source	SS	df	MS	F	p-value	Partial η²	Decision
Corrected Model	690.66	2	345.33	16.08	.000	.141	—
Intercept	1250.45	1	1250.45	58.25	.000	.229	—
Covariate (Pre-test)	92.44	1	92.44	4.31	.039	.022	—
Group (AI vs Non-AI)	598.22	1	598.22	27.89	.000	.125	Reject H_{01}
Error	4205.12	196	21.46	—	—	—	—
Corrected Total	**4895.78**	**198**	—	—	—	—	—

The ANCOVA results in Table 4 provide rigorous statistical confirmation of the observed differences in achievement. After adjusting for pre-test scores, the group effect is highly significant ($F = 27.89$, $p < .000$). This means that the difference in post-test performance between the AI-exposed and non-AI groups is not only large in magnitude, as seen in Table 1, but also statistically unlikely to be due to chance.

The significance of the covariate (pre-test scores) further indicates that students' prior knowledge had some influence on post-test achievement ($F = 4.31$, $p = .039$). However, the much larger F-value associated with the treatment effect demonstrates that exposure to AI learning tools had a far more substantial impact on achievement than prior knowledge. The ANCOVA finding strengthens the internal validity of the study in several ways:

1. It controls for baseline differences to isolate the true effect of treatment.
2. It confirms that the instructional method not initial ability explains the post-test variance.
3. It provides strong causal evidence that AI tools significantly enhance cognitive achievement.

In practical terms, this means that AI-supported instruction offers learning advantages beyond what is possible through traditional teaching alone. Table 4 therefore provides the most compelling evidence that the integration of AI tools into the teaching of Educational Statistics is not just beneficial but significantly transformative.

H_{02}: There is no significant difference in the mean achievement scores of male and female students exposed to AI tools.

Table 5: ANCOVA Summary for Difference in Achievement between Male and Female Students exposed to AI tools

Source	SS	df	MS	F	p-value	Partial η²	Decision
Corrected Model	99.65	2	49.83	2.40	.096	.047	—
Intercept	890.32	1	890.32	42.95	.000	.307	—
Covariate	85.37	1	85.37	4.12	.045	.041	—
Gender	14.28	1	14.28	0.69	.409	.007	Retain H_{02}
Error	2010.54	97	20.73	—	—	—	—
Corrected Total	**2110.19**	**99**	—	—	—	—	—

The ANCOVA results in Table 5 show that after adjusting for pre-test scores, there is no statistically significant difference in the post-test achievement scores of male and female students exposed to AI-based learning tools, $F(1,97) = 0.69$, $p = .409$. Although the covariate (pre-test score) was statistically significant, $F(1,97) = 4.12$, $p = .045$, indicating that prior knowledge slightly influenced post-test performance, the gender effect remained non-significant. This means that when baseline academic ability was controlled, male and female students benefited equally from AI-supported instruction. This finding reinforces the argument that AI-based learning tools create a gender-neutral and equitable learning environment. The adaptive feedback, private interaction interface, personalized explanations, and self-paced structure of AI tools likely reduced classroom anxiety and eliminated social learning barriers that sometimes create gender disparities in quantitative subjects. Therefore, $H0_2$ is retained. Gender does not significantly influence achievement among students exposed to AI-based learning tools.

$H0_3$: There is no significant difference in the mean achievement scores of male and female students not exposed to AI tools.

A second independent samples t-test was conducted.

Table 6: ANCOVA Summary for Difference in Achievement between Male and Female Students Not Exposed to AI Tools

Source	SS	df	MS	F	p-value	Partial η²	Decision
Corrected Model	100.07	2	50.04	2.52	.085	.049	—
Intercept	845.76	1	845.76	42.57	.000	.305	—
Covariate	78.62	1	78.62	3.94	.050	.039	—
Gender	21.45	1	21.45	1.08	.302	.011	Retain H_{03}
Error	1926.18	97	19.86	—	—	—	—
Corrected Total	**2026.25**	**99**	—	—	—	—	—

The ANCOVA results in Table 6 indicate that after adjusting for pre-test scores, there is no statistically significant difference in the post-test achievement scores of male and female students who were not exposed to AI tools, $F(1,97) = 1.08$, $p = .302$. The covariate effect is marginally significant, $F(1,97) = 3.94$, $p = .050$, suggesting that prior knowledge slightly influenced students' performance under traditional instruction. However, gender did not significantly affect achievement. This result shows that under conventional lecture-based instruction, male and female students performed similarly. However, their overall achievement levels remained lower compared to students exposed to AI tools, as earlier demonstrated. The finding implies that traditional instructional methods do not inherently favor one gender over another, but they may lack the adaptive and individualized support necessary to significantly elevate achievement for all learners.

Therefore, H_{03} is retained. There is no significant gender difference in achievement among students taught without AI tools.

Discussion of Findings

Students exposed to AI-based learning tools achieved higher post-test scores (M = 32.96) than those without AI exposure (M = 27.14), with the mean gain of the AI group (11.12) being more than double that of the control group (5.11). This finding demonstrates that AI tools significantly enhance students' mastery of Educational Statistics. The large mean difference and the significant ANCOVA result ($F = 27.89$, $p < .001$) indicate that the improvement is not due to chance but to the instructional treatment. This improvement can be attributed to several factors. First, AI-based tools provide immediate and continuous feedback, which helps students quickly identify and correct misconceptions before they become deeply rooted. Second, AI systems offer personalized learning pathways, allowing students to learn at their own pace and revisit difficult concepts such as probability distributions and hypothesis testing. Third, the interactive and simulation-based features of AI tools make abstract statistical concepts more concrete and easier to understand. Additionally, AI reduces dependency on one-size-fits-all instruction by providing adaptive difficulty levels, ensuring that both high- and low-performing students are adequately supported. These features collectively promote deeper cognitive processing, increased engagement, and better retention of knowledge.

This agrees with Holmes et al. (2021), who found that intelligent tutoring systems significantly improved students' learning outcomes in mathematics and statistics by providing adaptive feedback and individualized learning paths. Similarly, Ngugi & Eze (2020) reported that chatbots enhanced students' understanding of probability concepts due to real-time feedback.

Uche & Okoro (2023) also observed that students exposed to AI-powered simulations performed better in inferential statistics tests than those under traditional instruction. The finding also aligns with UNESCO (2022), which emphasized that AI tools improve conceptual clarity and support deeper learning through personalized interactions.

However, Olanrewaju (2022) cautioned that excessive dependence on AI may reduce independent problem-solving skills—implying that while AI boosts short-term performance, learners must still develop independent thinking habits.

Overall, the present study reinforces overwhelming global and local evidence that AI-supported learning enhances cognitive performance in quantitative subjects.

Male (M = 33.41) and female (M = 32.58) students exposed to AI tools recorded very similar mean scores. The difference was not statistically significant, $t(98) = 0.83$, $p = .408$. AI-based learning tools create a gender-neutral learning environment. Both genders received equal levels of adaptive support, personalized feedback, and learning flexibility. This finding may be explained by the nature of AI learning environments. AI tools provide a private and non-judgmental interface, allowing students to engage with learning materials without fear of criticism or embarrassment, which can sometimes affect participation in traditional classrooms. Additionally, AI eliminates teacher bias and classroom interaction differences, ensuring that all students receive equal instructional support regardless of gender. The self-paced learning structure also allows both male and female students to spend sufficient time mastering difficult concepts without pressure. Furthermore, AI systems respond to individual performance rather than gender, thereby promoting equity in learning outcomes.

This finding agrees with Rahman & Putri (2022), who reported that AI-based learning systems provide equitable learning opportunities for both genders because AI interfaces eliminate classroom bias. Yuan & Chiang (2021) also found no gender effect when students used automated feedback systems in statistics, noting that AI reduces performance anxiety common among females in quantitative subjects. Similarly, Mbata & Nwosu (2022) found that frequent engagement with AI practice tools improved achievement for both male and female students. The finding supports constructivist and Technology Acceptance Models (TAM), which emphasize that digital tools when well-designed neutralize social and psychological barriers to learning. This study adds evidence from the Nigerian context, showing that gender gaps typical of STEM subjects can be minimized when AI tools are integrated.

Male students (M = 27.70) scored slightly higher than females (M = 26.69), but the difference was not significant, $t(98) = 1.04$, $p = .302$. Traditional lecture-based teaching resulted in uniformly lower performance for both genders. The small difference suggests that both genders face similar challenges in teacher-centered environments such as limited feedback and lack of personalized support. This outcome may be due to the inherent limitations of conventional instructional methods. Traditional teaching approaches often rely on generalized instruction, which does not accommodate individual differences in learning pace and ability. Both male and female students are therefore equally affected by delayed feedback, making it difficult to promptly correct misunderstandings. In addition, the absence of interactive learning tools and simulations reduces students' ability to visualize and apply statistical concepts effectively. The lack of individualized attention in large classrooms further contributes to uniform learning difficulties across genders.

Moreover, factors such as statistics anxiety, poor mathematical background, and limited practice opportunities affect students irrespective of gender, leading to similar performance outcomes. Without adaptive learning support, both male and female students are constrained by the same structural limitations of traditional instruction.

This finding aligns with Nworgu & Ugwuanyi (2022) who found that students (male and female alike) struggle significantly with Educational Statistics due to inadequate practice opportunities. Nwafor & Onah (2022) similarly highlighted consistently low statistics achievement across genders in traditional classrooms. The result also agrees with Mensah (2021), who observed that students with limited digital literacy or no exposure to adaptive tools performed poorly regardless of gender.

Without AI tools, both male and female students are disadvantaged by the rigidity of traditional instruction. This reinforces the need for technology-enhanced teaching strategies to improve achievement and close persistent learning gaps.

Conclusion

The findings of this study provide compelling empirical evidence that Artificial Intelligence (AI)–based learning tools significantly enhance students' cognitive achievement in Educational Statistics. Students exposed to AI tools achieved markedly higher post-test scores ($M = 32.96$) than those taught without AI support ($M = 27.14$), and the ANCOVA test confirmed this difference as statistically significant, $F(1,196) = 27.89$, $p < .001$. The substantial mean gain advantage of the AI group (11.12) compared to the non-AI group (5.11) clearly demonstrates the instructional superiority of AI-supported learning environments.

Furthermore, the results showed no significant gender differences in achievement among students exposed to AI tools, $t(98) = 0.83$, $p = .408$. This indicates that AI learning systems provide an equitable, gender-neutral learning environment where both male and female learners benefit equally from adaptive feedback, personalized learning pathways, and private, non-judgmental interfaces. Similarly, no significant gender differences were found among non-AI students, $t(98) = 1.04$, $p = .302$, although both genders recorded lower overall performance compared to the AI-exposed group.

Overall, the study concludes that AI-based learning tools are not only effective but transformative in improving students' understanding and mastery of Educational Statistics. They bridge gaps inherent in traditional lecture-based instruction, provide individualized learning support, and reduce potential gender-based performance disparities. Therefore, the integration of AI tools into the teaching of Educational Statistics should be considered essential for modern, evidence-based, and equitable pedagogy in higher education.

Recommendations

Based on the findings and conclusion of this study, the following recommendations are made:

1. Universities should incorporate AI-driven learning platforms, intelligent tutoring systems, and adaptive statistical simulators into the regular instructional processes for Educational Statistics. This integration will promote deeper learning, improve achievement outcomes, and reduce difficulties commonly reported by students.
2. Lecturers should receive systematic training on effective use of AI tools for teaching, assessment, and feedback. Such capacity-building will ensure that instructors can confidently integrate AI into lesson delivery and maximize its benefits for students' cognitive achievement.
3. Institutional investment is required to provide stable internet connectivity, digital devices, AI-enabled computer laboratories, and licensed educational AI software. These facilities are essential for sustainable implementation of AI-supported instruction.
4. Students should be guided on how to use AI systems as learning enhancers rather than substitutes for personal effort. Structured orientation programmes should emphasize responsible use, academic

integrity, and strategies for combining AI guidance with independent problem-solving.

5. Educational Statistics syllabi should be revised to include activities that leverage AI feedback systems, automated practice drills, simulation-based tasks, and AI-enhanced assessments. These will support learner-centered pedagogy and continuous formative evaluation.

References

Adebayo, F., & Salami, A. (2023). *Utilization of artificial intelligence tools among Nigerian university students: Opportunities and challenges*. Journal of Digital Learning, 14(2), 55–68.

Anderson, L. W., & Krathwohl, D. R. (2021). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Chineke, P., & Oliver, C. (2022). *Awareness and utilization of AI-driven educational platforms among first-year undergraduates*. African Journal of ICT in Education, 7(1), 44–57.

Holmes, W., Bialik, M., & Fadel, C. (2021). *Artificial intelligence in education: Promises and implications*. Center for Curriculum Redesign.

Mbata, C., & Nwosu, G. (2022). *AI-based practice testing and academic performance among university students*. Nigerian Journal of Educational Technology, 5(2), 59–72.

Mensah, J. (2021). *Digital literacy and students' ability to benefit from AI-integrated instruction*. Journal of Educational Computing, 18(4), 233–247.

Ngugi, P., & Eze, B. (2020). Effect of chatbot-assisted learning on students' understanding of probability theory. *International Journal of STEM Education*, 9(3), 112–121.

Nwafor, U., & Onah, F. (2022). Challenges in teaching and learning educational statistics in Nigerian universities. *Nigerian Journal of Educational Research*, 17(1), 88–101.

Nworgu, B. G., & Ugwuanyi, L. (2022). Students' difficulties in learning educational statistics: Implications for curriculum planners. *Journal of Measurement and Evaluation*, 11(1), 25–38.

Olanrewaju, T. (2022). Consequences of overdependence on AI tutors in higher education. *West African Journal of Educational Research*, 4(2), 101–116.

Rahman, S., & Putri, D. (2022). Undergraduates' perceptions of AI-based learning systems in Indonesia. *Journal of Innovative Learning Technologies*, 7(3), 40–51.

Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Uche, A., & Okoro, P. (2023). Impact of AI-powered simulations on students' understanding of inferential statistics. *Contemporary Journal of ICT and Education*, 8(1), 77–92.

UNESCO. (2022). *Technology and artificial intelligence in education: Global trends and policy recommendations*. UNESCO Publishing.

Yuan, H., & Chiang, C. (2021). Automated feedback systems and learning outcomes in undergraduate statistics. *Asia-Pacific Journal of Education*, 41(2), 230–245.